Teaching Translation Technologies “Everyware”: Towards a Self-Discovery and Lifelong Learning Approach

Vanessa Enríquez Raído
The University of Auckland
v.enriquez@auckland.nz

ABSTRACT

This paper analyzes the main trends that have contributed to the reshaping of the current translation technology landscape, i.e. ubiquity, mobility, connectivity, and immediacy, and proposes a number of pedagogical reasons for promoting the teaching of translation technologies “everyware” as well as the transversal use of online information skills as means towards self-discovery and lifelong learning.

Keywords: Translation technology; ubiquity; mobility; connectivity; immediacy; computer-aided translation; “everyware” approach; self-discovery; lifelong learning; transversal skills; information skills; information literacy; multimedia literacy

RESUM (La docència de les tecnologies everyware de la traducció: un enfocament cap el creixement personal i el aprenentatge al llarg de la vida)

Aquest article analitza les tendències que han contribuït a reorientar el panorama actual de les tecnologies de la traducció, com ara la ubiqüitat, la mobilitat, la connectivitat i la immediatess, i proposa una sèrie de raons pedagògiques per tal de promoure la docència de les tecnologies de la traducció des de la perspectiva de la ubiqüitat tecnològica (“everyware”), així com la incorporació transversal de les destreses per a la gestió de la informació en línia com a recurs que afavoreix el creixement personal i l’aprenentatge al llarg de la vida.

Palabras clave: Tecnologías de la traducción; ubiqüitat; mobilitat; connectivitat; immediatess; traducción asistida per ordenador; enfocament transversal; creixement personal; aprenentatge al llarg de la vida; destreses transversals; destreses per a la gestió de la informació; alfabetització informàtica; alfabetització multimèdia

RESUMEN (La docencia de las tecnologías everyware de la traducción: un enfoque hacia el crecimiento personal y el aprendizaje a lo largo de la vida)

Este artículo analiza las tendencias que han contribuido a reorientar el panorama actual de las tecnologías de la traducción, como la ubicuidad, la movilidad, la conectividad y la inmediatez, y propone una serie de razones pedagógicas para promover la docencia de las tecnologías de la traducción desde la perspectiva de la ubicuidad tecnológica (“everyware”), así como la incorporación transversal de las destrezas para la gestión de la información en línea como recurso que favorece el crecimiento personal y el aprendizaje a lo largo de la vida.

Paraules clau: Tecnologías de la traducción; ubicuidad; movilidad; conectividad; inmediatez; traducción asistida por ordenador; enfoque transversal; crecimiento personal; aprendizaje a lo largo de la vida; destrezas transversales; destrezas para la gestión de la información; alfabetización informática; alfabetización multimedia
Introduction

With recent advances in translation automation technologies, above all of data-driven or statistical machine translation (MT), translation educational programs around the world are facing pressure from the market to offer up-to-date contents. Whether this pressure threatens to give priority to the (speedy) updating of curricula with modern subjects over the question of how these new contents might best be taught, is, and has always been, in my mind, very much subject to the various institutional, physical, socio-economic, and cultural contexts in which translator education is embedded. While it may be the case that some translation program managers and/or translator educators exert little effort to ground modern (and not-so-modern) subjects in state-of-the-art curriculum and pedagogical research, the growth in literature on the teaching and learning of recent subjects such as machine translation and post-editing, crowd translation, and cloud-based SaaS (Software as a Service) testify to a growing interest in curriculum development and course syllabi design.

Some of the most recent works in these areas include the proposals made by Doherty, Kenny, and Way (2012), Belam (2001), and O’Brien (2002) for designing courses on statistical machine translation (MT) and post-editing, Kelly and Way’s (2001) as well as Doherty and Moorkens’ (2013) pedagogical studies on the teaching of translation memory (TM) systems and MT, as well as Cánovas and Samson’s (2011), Mata Pastor’s (2008), and Flórez and Alcina’s (2011) suggestions for the use of open software in translator training, to name but a few. Other authors have contributed to this reconfiguration of training not so much in terms of making proposals for course contents but rather by identifying new skill sets and profiles (e.g. Rico and Torrejón, 2012) in what Anthony Pym has termed “a machine-translation age” (2012). Pym himself identifies a total of ten skills required for translation work using TM/MT technologies, which he arranges “under three heads: learning to learn, learning to trust and mistrust data, and learning to revise with enhanced attention to detail” (ibid.: unpaginated). I will be referring to these learning categories in more detail throughout this paper.

Other aspects that are typically associated with the teaching and learning of translation, including translation technology, relate to questions such as sequencing, assessment, pedagogical approaches, delivery methods, etc. The articles included in this journal issue, for instance, all refer to one or more of these themes. Ultimately, discussions on curriculum and syllabus design and development are influenced by the contextual factors mentioned above as well as by existing theories and models of translation and/or translator competence. Elisa Calvo (2011: 6), for example, states that competence models have been classified according to their theoretical background (e.g. cognitive, socio-constructivist, etc.); their experimental, observational or survey-based nature; the skills aimed at (e.g. those of an expert or a learner); their approach (e.g. multi-competence/maximalist vs. minimalist models); and purpose, i.e. whether the model “has been designed as a universal, all-purpose theory or as a more narrowly-focused model” (ibid.)

In line with Calvo’s paper on curriculum development, this paper does not aim at assessing “the quality or validity of the different competence models in place, nor to choose one as best for a specific curriculum” (2011: 7), in this case on translation technology. Again, the decision to adopt a particular type of curriculum over another is very much subject to the factors mentioned earlier as well as those summarized by Frank Austermühl in his contribution to this journal issue, i.e.: the fact that translation tools are developed outside universities, the lack of systematic research for designing course syllabi, the lack of technological expertise among some translator trainers, and time constraints within translation programs, among others. Rather, the purpose of this paper is to briefly describe some of the driving forces behind the current translation industry and their main implications for translator
training, where a plea is made for the adoption of an “everyware” (Greenfield 2006, see below) approach to the teaching of translation technologies and the transversal use of online information skills across the translation curriculum as a means towards self-discovery and lifelong learning.

1. Ubiquity

In 2005 I proposed a holistic approach to the teaching of translation technology by integrating it into practical translation courses, where the teaching situation would promote, at the same time, learning about, learning with, and learning within translation and technology environments (Enríquez-Raido, 2005, drawing on Torres del Rey, 2003). I started from the assumption that this neglect of didactic considerations was especially prominent with regard to the teaching of computer-aided translation (CAT). A cursory look at the curricula of Spanish and German translation programs at the time suggested that translation technology was primarily taught as an isolated and quasi translation-independent subject matter (cf. Kelly, 2005: 68). In other words, it seemed to be mainly dealt with as a separate, compartmentalized area of specialization within translation programs, taught as an end in itself rather than as a means towards the very end of translating, enacting knowledge and learning how to learn. I thus suggested that technology take a back seat to translation practice and become a means towards the end of translating rather than an end (or a course subject) in itself. I also stated that integrating technologies into translation practice courses would appear to dissolve what Watson (2001: 253) calls the “dichotomy of purpose” of vocational vs. pedagogical learning, thus helping to strengthen the relationship between otherwise fragmented areas of knowledge.

I am still a firm believer in the adoption of a “pedagogical holistic-bundle focus” (Kiraly, 2013: 199, drawing on Herold, 2010: 240; emphasis in the original) for the development of translation expertise in general and technological skills (see e.g. EMT Expert Group 2009: 7) in particular. Yet, my position today is that the teaching of translation technology should be relegated neither to a specific course on the subject nor to translation practice courses alone. Quite to the contrary, and very much in agreement with Pym (2012: unpaginated), the various technologies available to translators “should be used everywhere,” i.e. “as much as possible and in as many different courses as possible” (e.g. courses in documentary research, terminology management, project management, editing and revising for translators, and even translation theories, where different approaches to and methodologies of translation can be discussed against the current technological background). One of the main pedagogical reasons for promoting an integrated and transversal use of technologies across the translation curriculum is that, as Alcina, Soler, and Granell (2008: 231) as well as Pym (2012: unpaginated) point out, we are basically dealing with skills that are developed (at least initially) through imitation and repetition.

From an industry and business perspective, another reason for wanting to use technology “everyware”—to use Adam Greenfield’s (2006) more fitting term—in translator training relates to the ubiquity of digital content and information technology in what Greenfield has referred to as “The Dawning Age of Ubiquitous Computing” (ibid.). According to Greenfield, the concept of ‘everyware’ refers to the “ever more pervasive, ever harder to perceive computing [that] has leaped off the desktop and insinuated into everyday life” (ibid., cited in Shannon, 2010: 2). For Shannon, ubiquity implies that “digital content will continue to grow exponentially, as ‘good-enough’ methods of generating, delivering and translating content spread” (ibid.). With the industry “struggling to develop high-quality materials at low cost and high-volume” and where connecting the “Next Four Billion Users will necessitate more than 1,000 languages,” current translation quality practices such as translate-edit-proofread (TEP) “will not scale to such linguistic demands” (ibid.: 6). This new reality has led to a paradigmatic shift from
“translation as a relative expensive professional [activity] to translation as a utility” (Choudhury and McConnell, 2013: 52). Should translator training wish to reconcile with this paradigmatic change, our translation students should perhaps be exposed to the production of both high-quality and good-enough translations—at least as a way to raise consciousness on the challenge of ubiquity. Even though, as Shannon points out, “the translation community may at first bristle, good enough will need to become a part of language quality” (2010: 6). In fact, this has already been the case in the professional community with regard to translators’ attitudes towards the editing of different TM/MT matches. Pym (2012: unpaginated), for example, reports that “many of the experiments that compare TM/MT with fully human translation pick up a series of problems related to the ways translators evaluate the matches proposed to them,” including the inability to spot errors in the proposed matches, the editing of fuzzy matches that would have been better to translate from scratch, or the insufficient trust that translators place upon authoritative memories. The specific skills that translators would need to develop to “assess the trustworthiness of proposed matches” that are relevant to our discussion are:

2.1. Ability to check details of proposed matches in accordance with knowledge of provenance and/or the corresponding rates of pay (“discounts”). That is, if you are paid to check 100% matches, then you should do so; and if not, then not.

2.2. Ability to focus cognitive load on cost, beneficial matches. That is, if a proposed translation solution requires too many changes (probably a 70% match or below), then it should be abandoned quickly; if a proposed match requires just a few changes, then only those changes should be made; and if a 100% match is obligatory and you are not paid to check it, then it should not be thought about (ibid.).

One would suspect that translators’ decision to be either optimizers or satisficers (cf. Enríquez Raído 2011a: 19, drawing on Simon 1979, as well as Austermühl’s contribution to this journal issue) has to do with issues of quality and efficiency as much as with issues of payment and related working conditions.

2. Mobility and Immediacy

Other trends that have contributed to the reshaping of the industry landscape are mobility and immediacy (Shannon, 2010: 6). In terms of mobility, Choudhury and McConnell claim that with more and more users “shifting to laptop, tablet and mobile devices, most translation technology vendors have to develop a mobile strategy to remain competitive” (2013: 52). The authors of the “Translation Technology Landscape Report” further state that translation management systems, for example, should enable translators to work on projects from tablets and smartphones. Small form factor devices, of course, have limited real estate, so this is not a trivial problem. Companies that develop an intuitive mobile interface will have a distinct advantage in the marketplace compared to platforms that only work with conventional computers. This is particularly true for systems that enable crowd translation, since crowd translation often involves large numbers of people doing individually small amounts of work, a perfect use case for casual translation via a mobile device (ibid.).

Processes of crowd translation (also referred to as collaborative or volunteer translation in the literature) would ideally involve volunteers working alongside professional translators, “presumably in different phases of the workflow” (Pym, 2011: unpaginated; cf. Choudhury and McConnell, 2013: 51; Désilets 2010), so as to enable massive scale, low-cost translation while still maintaining quality standards.

Closely related to the notions of ubiquity and mobility, i.e. when information technology is ‘everyware’ and worldwide users are mobile, is that of “immediacy: the desire to consume information as quickly as possible” (Shannon, 2010: 11). Satisfying this desire requires real-time technology and multilingual communication to be within easy reach, a trend that, in turn,
has contributed to the convergence of technologies through the deployment of machine translation and its integration in translation memory systems as well as in other, more general applications and platforms used by the public-at-large. Here, Choudhury and McConnell (2013: 53) predict that successful CAT tools “will be connected to a variety of machine translation engines, usually via web services” (which is, in fact, already happening); that MT and speech recognition technology will be employed “in more business use scenarios in the near future, including healthcare and other sectors with high equity interactions;” that “machine translation, search tools and web content management systems will continue to converge, further embedding multilingualism on the internet and therefore across devices;” and that, as the quality of machine-translated output continues to improve, “machine translation will converge with consumer listening [i.e. understanding customers’ opinions about products, the topics of their discussions, where these discussions take place, etc., through social media monitoring, opinion mining, customer engagement, and competitor analytics] and a myriad of information processing technologies.”

The translation industry has mainly responded to issues of ubiquity, mobility, and immediacy by developing cloud-based SaaS, “coupled with real-time, multilingual communication engines” (Shannon, 2010: 2). From a translator training point of view, and particularly in terms of mobility and immediacy, it follows that students need to have network connectivity to access the various technologies from a wide range of devices (i.e. mobiles, laptops, tablets, desktops, etc., when and if they have access to these devices)—thus being able to do so from various operating systems (Windows, Mac, Linux, etc.)—and that they also have, where made available by the respective software developers, the editor versions of the various tools and applications installed on said devices to allow for offline translation in case of Internet connectivity failures. This, in turn, involves thinking about the various teaching spaces that would facilitate these technologized learning environments (cf. Pym 2012). In addition, and taking into account the issue of ubiquity as well, the current translation technology landscape implies developing multimedia literacy, understood here as the ability to critically use and evaluate various tools and applications (as well as to process different file formats) in a transversal way, i.e. not limited to a particular medium and, as pointed out above, across various translation-oriented activities. This would seem to be best achieved by having our students work with different (translation) technologies—again, in as many courses as possible—including proprietary tools and freeware, and critically assess which ones suit which translation projects and/or specific tasks best. In this regard, there seems to exist a consensus within the teaching community that students “should not learn just one tool step-by-step” (Pym 2012: unpaginated). Well, certainly not every single MT/TM or CAT tool. Thus, I think that the focus should gradually move towards the understanding of core functions and concepts rather than following specific steps (which may very well suit the early stages of translator training). Once students acquire solid declarative as well as procedural knowledge about a specific tool or application, they will most likely be in a strong position to transfer said knowledge and skills to other implementations of the same tool type and working scenarios. After all, most CAT and MT tools share the same functionality. This also calls for a self-discovery approach—which obviously requires the right nurturing environment (mainly through scaffolding)—where students “have to be left to their own devices, as much as possible, so they can experiment and become adept at picking up a new tool very quickly, rely on intuition, peer support, online help groups, online tutorials, instruction manuals, and occasionally a human instructor to hold their hand when they enter panic mode” (ibid.) or

---

1 Naturally, this would require students to invest time outside lecture/tutorial/lab hours (at the University of Auckland, for example, students are expected to work 20 hours per week for a 30-point course based on three face-to-face contact hours per week) to experiment with various software applications. This presupposes, in turn, that they have easy access to these applications. At my institution, for instance, we have two computer labs that provide said access to students. Yet, they much rather prefer installing the various required applications (either as demo versions or freeware) on their own laptops and having the flexibility to work at their own pace from any location they choose.
simply when they need guidance on whichever level of inquiry. Here, collaborative and socio-constructivist approaches have also proved successful epistemologies for the teaching of translation technology.

My own postgraduate students, for example, recently prepared a number of workshops on various second-generation (i.e. cloud-based) CAT tools, such as Wordbee, XTM Cloud, MemSource Cloud, and TM Anywhere) based on the knowledge and skills they acquired just by learning one specific first-generation (i.e. client/server based) CAT tool and extrapolating from there. This naturally required them to do all the things mentioned above, i.e. employing a wide range of resources for self-learning purposes (more on learning how to learn below), experimenting with the tools (through trial and error) and resorting to peer and/or teacher support when needed. I learned a great deal of things by participating in the workshops. In fact, I could not have done a better job myself, certainly not with the same degree of thoroughness and critical stance as that of my students, who probably do not have the same high workload that I have (or maybe they do and I just need to prepare my classes better).

For example, I learned that the advantages of second-generation over first-generation CAT tools can be numerous, such as accessibility across a wide range of devices and operating systems (i.e. cloud-based SaaS is not tied to the Windows OS), including mobile devices; their subscription-based model vs. the high costs associated with the payment of permanent licenses; improved software release cycles and upgrades; improved user interfaces; high capacity storage (of very large memories or LVMs), etc. I also learned that second-generation CAT tools need not render the use of client/server based applications obsolete, at least not for now, as the latter can be upgraded to interoperate with new cloud-based translation systems. But, most of all, I learned, yet again, that having my students engage in self-discovery and self-learning at the appropriate level of development is more beneficial to them than having me demonstrate one, two or three tools thoroughly (or at least, so they tell me).

In terms of self-discovery, we find a number of really interesting pedagogical proposals coming from the field of process-oriented translation studies, where the tools used for research purposes become learning tools in themselves. One of the most direct ways of doing so is to have students engage in actual process research in the classroom, “both as a means of self-discovery and as an approach to learning about research” (Pym, 2009: 136). Pym presents this mode of application by using screen recording software like TechSmith’s Camtasia Studio or Blueberry’s BB FlashBack as a direct training tool to conduct three experiments in the classroom (one on machine translation, another on translator styles, and a third on time pressure and its effects). Verbal and written protocols have also been widely used not only for data-elicitation but also as pedagogical tools to increase students’ awareness about their translation problem-solving and decision-making behaviors. Others, like Fabio Alves (2005), report on the direct application of the keystroke-logging program Translog. Pekka Kujamäki (2010), similar to Pym, reports on the direct use of screen recording as a training tool for students to reflect upon their own translation processes as well as those of fellow students, while Erik Angelone (2013) discusses the pedagogical advantages of also directly using screen recording as a self-training tool for students to model expert translation processes. In addition, my own pedagogical research (Enríquez Raído, 2013a), like that of Kujamäki (2010) and Massey and Ehrensberger-Dow (2011a), shows that screen recording can also be a useful teacher’s diagnostic tool, at least in beginner classes. This particular use of screen recording may help teachers direct students’ attention to processing aspects that may otherwise go unnoticed. Self-discovery, finally, requires the ability of learning how to learn, which I have previously referred to as information and/or web search literacy (cf. Enríquez Raído, 2011a, 2011b, 2012, 2013b, and which I discuss in more detail below with regard to the use of various data sources and information resources, and the need to critically assess these for translation purposes.
3. Big Data

Another significant driver behind translation technology advances is the growing use of large amounts of data, or what Choudhury and McConnell (2013: 52) refer to as “big data”. Linguistic data (e.g. speech and text corpora, glossaries, lexicons, and grammars) is not only the main driver for language and translation technologies, it is also “used to train and enhance the quality of the output generated by such technologies (ibid.). In recent years, we have witnessed a shift of focus from translation memories that are stored on local hard disks to massive amounts of translation data stored in the cloud, mainly in the form of parallel text corpora. The more this data is shared and openly accessible—publicly accessible translation memories include those made available by the European Commission’s Directorate-General for Translation (DGT), the Translation Automation User Society (TAUS), MyMemory, Linguee, and Glosbee, among others (see http://revistes.uab.cat/tradumatica/index.php/Publicly_accessible_translation_memories_(TMs) for details)—the more it will be used by a wide range of users and, in principle (see below), the better data-driven translation systems will perform. This seems to be particularly true for data that is curated, i.e. standardized in terms of file formats (e.g. to allow for machine readability) and categorized into ontologies, as well as pre-processed “through tokenization or part of speech tagging for example” (ibid.).

As Choudhury and McConnell remark, the best sources for accumulating “translation data are good quality human translations from trusted sources, such as government bodies and institutions, companies large and small, from professional translators and consumers themselves” (ibid.). Another way of aggregating sizeable amounts of data is crawling from the World Wide Web. Yet, the risk of obtaining less quality data in this way is considerably higher than obtaining data from more trusted sources like the ones mentioned above. Regardless of the method employed for data aggregation, and as Pym (2012: unpaginated) points out regarding the use of machine translation, when free MT/TM technologies become ubiquitous, the risk of “recycling errors that are fed back into the very databases on which” these technologies operate can be significant, thus causing quality-related problems for translation purposes. There are various ways of coping with quality management issues from a technical perspective, such as automatically cleaning and improving the quality of source and target texts as well as customizing machine translation engines (Choudhury and McConnell: 60-61). From a translator training point of view, having to deal with large amounts of data available in various forms (e.g. parallel text corpora, terminology databases, dictionaries, glossaries, etc.) has strengthened the need to emphasize two important aspects. The first concerns the shift of focus from data generation to data selection. As Pym remarks,

whereas much of the translator’s skill set and effort was previously invested in identifying possible solutions to translation problems (i.e. the “generative” side of the cognitive process), the vast majority of those skills and efforts are now invested in selecting between available solutions. And then adapting the selected solution to target, side purposes (i.e. the “selective” side of the cognitive processes). [...] That is a very simple and quite profound shift, and it has been occurring progressively with the impact of the Internet (2012: unpaginated; cf. Pym 2003).

Indeed, the impact and penetration of the Internet have dramatically transformed the way translators carry out their research. Among others, the appearance of the Internet has eliminated previous constraints of time and space regarding the acquisition of information. Yet, its ubiquity and structure along with the dispersed and dynamic nature of the information available on the Web continue to pose a number of challenges regarding the critical evaluation, selection, and use of credible sources of information. As Pinto Molina and Sales Salvador also remark, “[f]rom the user’s viewpoint, what matters is that today the problem is not so much accessing as locating the information needed to resolve a concrete problem or take a particular decision relating to one’s work or daily life, realizing the needs of quality, rigour, suitability, depth and specificity” (2007: 534). This also applies to the evaluation of
matches/proposals found in the various data sources available within a particular translation environment tool (TEnT), where discrepancies within and across translation memories, machine-translated content and term bases, for example, require translators to critically assess the provenance and reliability of the information provided in said databases. Oftentimes, it may also require cross-checking said information with external sources of consultation.

Although data sources in non-digital format are available to, and are indeed widely used by translators, today the Web is probably the resource most frequently used by these professionals. Here, De Schryver’s (2002) distinction of the use of the Web for Corpus—in which the Web “is used as a source of texts in digital format for the subsequent implementation of an offline corpus” (Buendía Castro and López Rodríguez, 2013: 54)—and the Web as Corpus—which uses the Web directly as one mega-corpus from which data is retrieved through direct search engine queries—becomes important for translators and other language-related professionals (see Austermühl, 2012 and his contribution to this journal issue for a more detailed discussion of De Schryver’s distinction). The latter approach has been the focus of my own research for the development and integration of information literacy (cf. Massey and Ehrensberger-Dow (2011b), in particular of web search skills in translator training (Enríquez Raído, 2011a, 2013b). Here, I have also proposed the transversal use of online information search and evaluation skills not only as a quasi-universal means to problem solving and decision making (regardless of any specialised domain) but also as a means to learn how to learn, in this case about the various technologies and online resources available to translators. The ability to develop online information skills is in fact the second important aspect I referred to above regarding the main consequences of having to deal with big data.

To sum up, the pedagogical guidelines proposed here do not, as Pym (2012: unpaginated) puts it, “operate through fixed recipes”. Rather, these guidelines are based on key implications drawn from the main trends that have contributed to the reshaping of the current translation technology landscape and that foster a transversal, i.e. an “everyware” approach to the teaching of translation technologies, favoured by processes of self-discovery and lifelong learning.

References


2 Information literacy, as defined by the American Library Association, involves “a set of abilities requiring individuals to recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information” (ALA 1989, cited in Association of College and Research Libraries 2000: 2).


———. (2012). “Information Literacy in Translator Training: Research Opportunities and Didactic Implications”. The 1st International Conference on Research into the Didactics of Translation (DidTrad)/ VIII International conference on Translation, Universitat Autonoma de Barcelona, Barcelona, Spain, 21 June - 22 June 2012.
“Everyware”: Towards a Self-Discovery and Lifelong Learning Approach
Vanessa Enríquez-Raído


http://revistes.uab.cat/tradumatica
Els continguts de la revista estan subjectes a una llicència Creative Commons (CC BY 3.0)


